Chap. 7. Properties of expectations.

§ 7.1 Introduction.

Recall that in the discrete case

$$E[X] = \sum_x x\, p(x).$$

In the cts case,

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$$

The expectation of X is a weighted average of the possible values of X.

§7.2 Expectation of functions of r.v.'s and sums of r.v.'s

Prop 2. Let $g : \mathbb{R}^2 \to \mathbb{R}$.

(1) If both X and Y are discrete with a joint prob. mass function $p(x, y)$, then

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)\, p(x,y)$$

(2) If X, Y are jointly cts with a density $f(x,y)$, then

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy$$

Pf. It is similar to the one r.v. case. $\square$

Corollary 3. $E[X+Y] = E[X] + E[Y]$.

Pf. We only prove the result in the case when

X, Y are jointly cts.

Then by Prop. 2, $E[X+Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f(x,y)\,dx\,dy$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\,dx\,dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x f(x,y)\,dy\right)dx + \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} y f(x,y)\,dx\right)dy$$

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx + \int_{-\infty}^{\infty} y f_Y(y)\,dy$$

$$= E[X] + E[Y]. \quad \boxtimes$$

By induction, we have the following

Corollary 4.

$$E[X_1 + \cdots + X_n] = \sum_{i=1}^{n} E[X_i].$$

§ 7.4 Covariance.

Recall the variance of a r.v. X is given by

$$\mathrm{Var}(X) = E[(X-\mu)^2], \text{ where } \mu = E[X].$$

It describes how far is X from its mean.

Def. (Covariance)

Let X, Y be two r.v.'s. The covariance of X and Y, denoted by $\mathrm{Cov}(X,Y)$, is defined by

$$\mathrm{Cov}(X,Y) = E[(X-E[X])\cdot(Y-E[Y])].$$

In particular, $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$.

Lem 4. Let X, Y be independent, and $g, h: \mathbb{R} \to \mathbb{R}$.

Then $$E[g(X)\cdot h(Y)] = E[g(X)]\cdot E[h(Y)].$$

pf. We only prove it in the cts case.

$$E[g(X)h(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)\,h(y)\,f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)\, h(y)\, f_X(x)\, f_Y(y)\, dx\, dy$$

$$= \left(\int_{-\infty}^{\infty} g(x) f_X(x)\, dx\right)\left(\int_{-\infty}^{\infty} h(y) f_Y(y)\, dy\right)$$

$$= E[g(X)] \cdot E[h(Y)].$$

Corollary 5. If X, Y are independent, then $\mathrm{Cov}(X, Y) = 0$.

pf. When X, Y are independent, by Lem 4,

$$\mathrm{Cov}(X,Y) = E\left[(X - E[X])(Y - E[Y])\right]$$

$$= E\left[(X - E[X])\right] \cdot E\left[(Y - E[Y])\right]$$

$$= 0$$

$\square$

Remark: $Cov(X, Y) = 0$ does not imply that X, Y are independent.

Example 6. Let X, Y be two r.v.'s such that

① $P\{X=0\} = P\{X=-1\} = P\{X=1\} = \frac{1}{3}$

② $Y = \begin{cases} 0 & \text{if } X \neq 0. \\ 1 & \text{if } X = 0. \end{cases}$

- A short cut formula

$$Cov(X, Y) = E[XY] - E[X]E[Y].$$

- $E[X] = \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot (-1) + \frac{1}{3}(1) = 0$

$P\{XY = 0\} = 1 \Rightarrow E[XY] = 0 \cdot 1 = 0$

Hence $\mathrm{Cov}(X, Y) = 0$.

But X, Y are not independent.

$$P\{X=0, Y=0\} = 0$$

But $P\{X=0\} = \frac{1}{3}$

$$P\{Y=0\} = P\{X \neq 0\}$$

$$= P\{X=1\} + P\{X=-1\}$$

$$= 2/3$$

Hence $P\{X=0, Y=0\} \neq P\{X=0\} \cdot P\{Y=0\}$

Therefore, X and Y are not independent.

Prop 1. (1) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.

(2) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

(3) $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$, $a \in \mathbb{R}$.

(4) $$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}(X_i, Y_j).$$

(1), (3), (4) imply that $\mathrm{Cov}(\cdot, \cdot)$ is bi-linear.

Pf. Let us prove (4) only. Write $\mu_i = E[X_i]$,

$\nu_i = E[Y_j]$.

Then by definition,

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = E\left[\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right)\cdot\left(\sum_{j=1}^{m} Y_j - \sum_{j=1}^{m} \nu_j\right)\right]$$

$$= E\left[\sum_{i=1}^{n}(X_i - \mu_i) \sum_{j=1}^{m}(Y_j - \nu_j)\right]$$

$$= E\left[\sum_{i=1}^{n}\sum_{j=1}^{m}(X_i - \mu_i)(Y_j - \nu_j)\right]$$

By the linearity of E

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} E\left[(X_i - \mu_i)(Y_j - \nu_j)\right]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j).$$

$\square$

Corollary 2. $$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\cdot\sum_{1\le i<j\le n} Cov(X_i, X_j).$$

Moreover if $X_1, \cdots, X_n$ are pairwise independent,

then

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i).$$

§ 7.5 Conditional expectations.

Def. If X and Y are discrete, then the conditional expectation of X given $Y=y$, is

$$E[X|Y=y] := \sum_x x\, P\{X=x | Y=y\}$$

provided that $P\{Y=y\} > 0$.

Def. In the case when X and Y are jointly cts with a density $f(x,y)$, the conditional expectation of X given $Y=y$, is defined by

$$E[X|Y=y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)\, dx,$$

provided that $f_Y(y) > 0$, where

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}.$$

Example 1. Let X, Y be jointly cts with a density

$$f(x,y) = \begin{cases} e^{-x/y} \cdot e^{-y}/y & \text{if } x, y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Calculate $E[X \mid Y=y]$, $y>0$.

Solution: $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$

$$= \int_0^{\infty} e^{-x/y} e^{-y}/y \, dx$$

$$= -e^{-x/y} e^{-y} \Big|_{x=0}^{\infty}$$

$$= e^{-y}, \quad \text{if } y>0$$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$= e^{-x/y}/y \quad \text{if } x, y > 0.$$

$$E[X|Y=y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, dx$$

$$= \int_0^{\infty} x \cdot e^{-x/y} / y \; dx$$

$$\overset{\text{Int by Part}}{=} x \cdot \left(-e^{-x/y}\right)\Big|_{x=0}^{+\infty} + \int_0^{\infty} e^{-x/y} \cdot dx$$

$$= 0 + \left(-y e^{-x/y}\right)\Big|_{x=0}^{+\infty}$$

$$= y \quad \text{if} \quad y > 0.$$

$\blacksquare$

Now write

$E[X|Y]$ as a function of Y by

$$y \longmapsto E[X|Y=y]$$

$E[X|Y]$ is a r.v., the value of which depends on the value of Y.

Prop 2. $E[X] = E[\,E[X|Y]\,]$

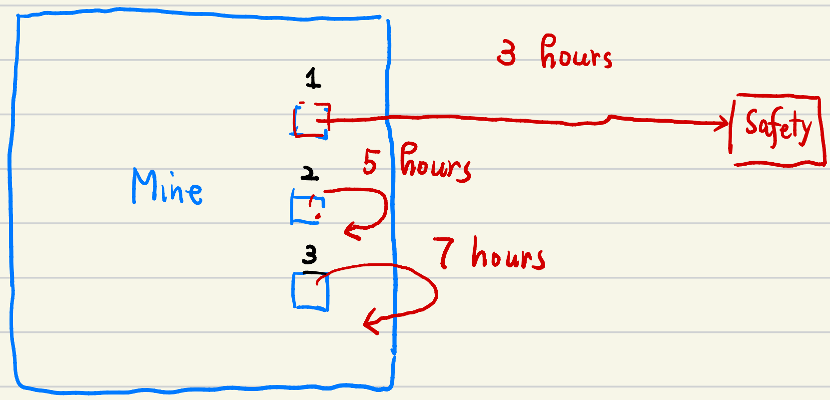
Pf. We only prove it in the discrete case.

$$\begin{aligned}
E[\,E[X|Y]\,] &= \sum_y E[X|Y=y] \cdot P_Y(y) \\
&= \sum_y \cdot \sum_x x \cdot P\{X=x|Y=y\} \cdot P_Y(y) \\
&= \sum_y \sum_x x\, P\{X=x, Y=y\} \\
&= \sum_x \sum_y x\, P\{X=x, Y=y\} \\
&= \sum_x x\, P\{X=x\} = E[X]
\end{aligned}$$

$\square$

Example 3.

A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hours of travel. The second door leads to a tunnel that will return him to the mine after 5 hours of travel. The third door leads to a tunnel that will return him to the mine after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?



Solution: Let X denote the length of time (in hours) until the miner reaches safety.

Let Y denote the door that he choose in the first time.

By Prop 2,

$$E[X] = E[E[X|Y]]$$

$$= E[X|Y=1]\cdot P\{Y=1\}$$

$$+ E[X|Y=2]\cdot P\{Y=2\}$$

$$+ E[X \mid Y=3] \cdot P\{Y=3\}$$

$$= \frac{1}{3}\Big(E[X \mid Y=1] + E[X \mid Y=2] + E[X \mid Y=3] \Big)$$

$$= \frac{1}{3}\Big(3 + (5 + E[X]) + (7 + E[X]) \Big)$$

Solving this equation, we obtain

$$E[X] = 3 + 5 + 7 = 15 \quad \text{(hours)}.$$

$\blacksquare$

§ 7.7. Moment generating functions.

Def. Let X be a r.v. and $t \in \mathbb{R}$. Define

$$M(t) = E[e^{tX}]$$

and we call $M(t)$ the moment generating function of X

Remark: • Since

$$e^{tX} = \sum_{n=0}^{\infty} \frac{t^n}{n!} X^n ,$$

we have

$$M(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E[X^n] \qquad ①$$

- $M(0) = 1$.
- For some random variables X, $M(t)$ may not exist for some $t \in \mathbb{R}$.

For instance

$$P\{X = n\} = \frac{1}{n(n+1)}, \quad n = 1, 2, \dots$$

Then $E[X^k] = \infty \quad \forall\, k = 1, 2, \dots,$ so

$$E(e^{tX}) = \infty \quad \forall\ t > 0.$$

Prop 1. Suppose $M(t)$ exists and is finite on a neighborhood $(-t_0, t_0)$ of 0, then

$$M^{(n)}(0) = E[X^n], \quad n = 1, 2, \dots$$

Example 2. Let X be a binomial r.v. with parameters (n, p). Calculate $M(t)$.

$$M(t) = E[e^{tX}] = \sum_{k=0}^{n} e^{tk} \cdot \binom{n}{k} p^k \cdot (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (e^t p)^k (1-p)^{n-k}$$

$$= (e^t p + (1-p))^k$$

Example 3. Let X be the Poisson r.v. with parameter λ. Calculate $M(t)$.

Solution.

$$M(t) = E[e^{tX}]$$

$$= \sum_{n=0}^{\infty} e^{tn} \cdot \frac{\lambda^n}{n!} e^{-\lambda}$$

$$= \sum_{n=0}^{\infty} \frac{(e^t\lambda)^n}{n!} e^{-\lambda}$$

$$= e^{e^t\lambda} \cdot e^{-\lambda}$$

$$= e^{\lambda(e^t-1)}.$$

Example 4. Let X be a standard normal r.v. Calculate $M(t)$.

Solution:

$$M(t) = E[e^{tX}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} \cdot e^{\frac{-x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \cdot \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}} dx$$

$$= e^{t^2/2}.$$

Example 5. Let X be a normal r.v. with mean μ and Variance.

Calculate $M(t)$ for X.

Solution: Let $Z = \frac{X-\mu}{\sigma}$. Then Z is a standard normal r.v.

Hence

$$M(t) = E[e^{tX}] = E[e^{t(\mu+\sigma Z)}]$$

$$= e^{t\mu} \cdot E[e^{t\sigma Z}]$$

$$= e^{t\mu} \cdot e^{t^2\sigma^2/2}$$

$$= e^{\frac{t^2\sigma^2}{2} + t\mu}.$$

Prop 6. If X, Y are independent, then

$$M_{X+Y}(t) = M_X(t)\, M_Y(t).$$

Pf.

$$M_{X+Y}(t) = E[e^{tX+tY}]$$

$$= E[e^{tX} \cdot e^{tY}]$$

$$= E[e^{tX}] \cdot E[e^{tY}].$$

$$= M_X(t)\, M_Y(t)$$